UNIVERSIDAD CARLOS III DE MADRID

working papers

Working Paper 12-21
Statistics and Econometrics Series 15
Julio 2012

Departamento de Estadística
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624-98-49

A VECTOR OF DIRICHLET PROCESSES

Fabrizio Leisen [1], Antonio Lijoi [2], Dario Spanó [3]

Abstract

Random probability vectors are of great interest especially in view of their application to statistical inference. Indeed, they can be used for determining the de Finetti mixing measure in the representation of the law of a partially exchangeable array of random elements taking values in a separable and complete metric space. In this paper we describe a construction of a vector of Dirichlet processes based on the normalization of completely random measures that are jointly infinitely divisible. After deducing the form of the Laplace exponent of the vector of the gamma completely random measures, we study some of their distributional properties. Our attention particularly focuses on the dependence structure and the specific partition probability function induced by the proposed vector.

Keywords: Bayesian inference, Dirichlet process, Gauss hypergeometric function, Multivariate Levy measure, Partial exchangeability, Partition probability function

[1] Fabrizio Leisen, Departamento de Estadística, Universidad Carlos III de Madrid, C/ Madrid 126, 28903 Getafe (Madrid), España, e-mail: fabrizio.leisen@uc3m.es

[2] Antonio Lioji, Department of Economics and Business, University of Pavia, via san felice 5, 27100 Pavia, Italy. Also affiliated to Collegio Carlo Alberto, Moncalieri (TO), Italy, email: lijoi@unipv.it

[3] Dario Spanó, Departament of Statistics, University of Warwick, Coventry CV4 7AL, United Kingdom, e-mail: d.spano@warwick.ac.uk

1. Introduction

Random probability measures represent a key ingredient for the actual implementation of Bayesian nonparametric procedures and the Dirichlet process is the first example that has appeared in the literature. See [7]. The distribution of a random probability measure is typically used as the de Finetti measure of an infinite exchangeable sequence of random elements $(X_n)_{n\geq 1}$ taking values in some complete and separable metric space $\mathbb{X}$ endowed with the Borel σ–algebra $\mathscr{X}$. In other terms, if $\boldsymbol{P}_{\mathbb{X}}$ stands for the set of all probability measures on $(\mathbb{X}, \mathscr{X})$ and $\mathscr{P}_{\mathbb{X}}$ is the σ–algebra induced by the topology of weak convergence on $\boldsymbol{P}_{\mathbb{X}}$, one has

$$\mathbb{P}[(X_1,\dots,X_n)\in A] = \int_{\boldsymbol{P}_{\mathbb{X}}} p^n(A)\, Q(\mathrm{d}p) \qquad \forall A\in\mathscr{X}^n, \quad \forall n\geq 1 \tag{1}$$

where $p^n = \prod_1^n p$ and Q is a probability measure on $(\boldsymbol{P}_{\mathbb{X}}, \mathscr{P}_{\mathbb{X}})$. One of the most popular applications concerns nonparametric hierarchical mixture models for density estimation where the X_n's are latent variables and Q is the law of a Dirichlet process. See [18]. The dramatic advances in the implementation of Markov Chain Monte Carlo simulation algorithms in the last two decades have, then, made Bayesian nonparametric methods directly applicable to a wide range of real world problems. Moreover, a considerable body of work has been devoted to the proposal of alternatives to the Dirichlet process, i.e. different Q's in (1), for modelling exchangeable data. See [17] for a recent review.

Motivated by applications to regression problems, there has recently been great interest in the definition of nonparametric priors that accommodate for forms of dependence more general than exchangeability. In this case, instead of a single random probability measure $\tilde{P}$, one has a collection $\{\tilde{P}_z : z\in\mathcal{Z}\}$ of possibly dependent random probabilities indexed by a set of covariates z. Moreover, instead of a sequence of exchangeable random elements $(X_i)_{i\geq 1}$, one has an array $\{(X_i(z))_{i\geq 1} : \ z\in\mathcal{Z}\}$ such that for any $q\geq 1$, positive integers $n_1,\dots,n_q$ and $A\in\mathscr{X}^{|\boldsymbol{n}|}$

$$\mathbb{P}[(\boldsymbol{X}^{n_1}(z_1),\dots,\boldsymbol{X}^{n_q}(z_q))\in A] = \int_{\boldsymbol{P}_{\mathbb{X}}^{|\boldsymbol{n}|}} (p_{z_1}^{n_1}\times\cdots\times p_{z_q}^{n_q})(A)\, Q_{\mathcal{Z}}(\mathrm{d}p_{z_1},\dots,\mathrm{d}p_{z_q}) \tag{2}$$

where $\boldsymbol{X}^{n_j}(z_j) = (X_1(z_1),\dots,X_{n_j}(z_j))$, $\boldsymbol{n}=(n_1,\dots,n_q)$ and $|\boldsymbol{n}| = n_1+\cdots+n_q$. The mixture representation in (2) characterizes a *partially exchangeable* collection of random elements $\{(X_n(z))_{n\geq 1} : \ z\in\mathcal{Z}\}$. A widely used approach for defining a distribution $Q_{\mathcal{Z}}$ in (2) consists in fixing the (marginal) distribution of each $\tilde{P}_z$ and, then, suitably defining some form of dependence among different $\tilde{P}_z$'s. For example, when each $\tilde{P}_z$ can be seen as the normalization of a completely random measure (CRM) $\tilde{\mu}_z$, then dependence between any two random probabilities can be induced by dependence between the corresponding CRM's. This is pursued, for example, in [6], [14] and in [16]. This very same approach is undertaken henceforth and will lead to the definition of a new vector of dependent Dirichlet processes.

Under this approach it is usually hard to provide an analytical description for the dependence structure of the model, especially for vectors with more than two coordinates and with a non-Markovian type of dependence. The aim of this paper is to introduce a class of vectors of dependent gamma random measures, of arbitrary dimension, whose coordinates are exchangeable (thus non-Markovian) and whose dependence structure is analytically tractable. Moreover, vectors in our class have joint independent increments governed by a Levy copula with an Archimedean type of symmetry.

Another convenient (and popular) strategy for defining $Q_{\mathcal{Z}}$ applies when $\tilde{P}_z$, for any z in $\mathcal{Z}$, admits a stick–breaking representation. This means that, for any z in $\mathcal{Z}$, $\tilde{P}_z \stackrel{d}{=} \sum_{j\geq 0} \tilde{\pi}_{j,z}\, \delta_{Y_{j,z}}$ where the $\pi_{j,z}$ are determined via a stick–breaking construction and the $Y_{j,z}$ are iid from some distribution P_0. Hence, dependence between any two $\tilde{P}_z$ and $\tilde{P}_{z'}$, for any $z \neq z'$, is induced by specifying some dependence between $\pi_{j,z}$ and $\pi_{j,z'}$ or between $Y_{j,z}$ and $Y_{j,z'}$. This is the main idea inspiring the work of [19]. Recent contribution to this area are very well summarized in the review papers by [5] and [27]. Such a constructive approach has usually the advantage of making it possible to describe the dependence relationship in a context where it is hard or impossible to gain an analytical insight on the joint distribution of the vector. The class of dependent random measures treated in this paper may serve as a viable, complementary alternative to the constructive approach to defining dependent non-parametric prior models.
The outline of the paper is as follows. In Section 2 we shall introduce some basic elements on CRMs whereas Section 3 a new vector of dependent gamma CRMs will be introduced. We will point out some characteristic aspects of the dependence it induces. This vector of dependent CRMs is then normalized to yield a collection of dependent Dirichlet processes which correspond to a covariate space $\mathcal{Z} = \{1, \ldots, M\}$, for some positive integer M. In Section 4 we finally discuss some distributional results featured by the new process: in particular we shall point out a hint on the partition structure it gives rise to.

2. Some preliminaries

Among the different possibilities emerged in the literature for defining a probability distribution Q on $(\boldsymbol{P}_{\mathbb{X}}, \mathscr{P}_{\mathbb{X}})$, we shall focus on a strategy that makes use of completely random measures. Indeed, denote by $\boldsymbol{M}_{\mathbb{X}}$ that space of boundedly finite measures on $(\mathbb{X}, \mathscr{X})$ endowed with the so–called weak$^\sharp$ topology (see [3] for details) and let $\mathscr{M}_{\mathbb{X}}$ stand for the corresponding Borel σ–algebra. A random element $\tilde{\mu}$ defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ taking values in $(\boldsymbol{M}_{\mathbb{X}}, \mathscr{M}_{\mathbb{X}})$ and such that, for any pair of disjoint sets A and B in $\mathscr{X}$, the random variables $\tilde{\mu}(A)$ and $\tilde{\mu}(B)$ are independent. is a *completely random measure* (CRM). If it is assumed, as we shall do henceforth that $\tilde{\mu}$ does not have random masses at fixed locations, then

$$\tilde{\mu} = \sum_{i \geq 1} J_i \, \delta_{X_i}$$

where δ_c is the unit mass at point c and $\{(J_i, X_i) \, : \, i = 1, 2, \ldots\}$ are the points of a Poisson process on $\mathbb{R}^+ \times \mathbb{X}$ with intensity measure ν such that

$$\int_{\mathbb{R}^+ \times B} \min\{s, 1\} \, \nu(\mathrm{d}s, \mathrm{d}x) < \infty \qquad \forall B \in \mathscr{X}$$

The measure ν is also referred to as the intensity of $\tilde{\mu}$ itself. If it is further assumed that $\nu(\mathbb{R}^{\times}\mathbb{X}) = \infty$, one can show that $\tilde{\mu}(\mathbb{X}) \in (0, \infty)$, a.s., and it is possible to set Q in (1) as the probability distribution of $\tilde{p} = \tilde{\mu}/\tilde{\mu}(\mathbb{X})$. See [24]. This possible approach is already pointed out in [7] where it is shown that the Dirichlet process can be defined as the normalization of a gamma CRM, namely a CRM with intensity $\nu(\mathrm{d}s, \mathrm{d}x) = s^{-1} \, \mathrm{e}^{-s} \mathrm{d}s \, H(\mathrm{d}x)$ for some measure H on $\mathbb{X}$. Analogously, in [12] the normalized σ–stable CRM is introduced and it coincides with the normalization of a CRM with intensity $\nu(\mathrm{d}s, \mathrm{d}x) = \sigma s^{-1-\sigma} \, \mathrm{d}\sigma \, H(\mathrm{d}x)/\Gamma(1-\sigma)$, where $\sigma \in (0, 1)$. Also recall that for any measurable function $f : \mathbb{X} \to \mathbb{R}$ such that $\tilde{\mu}(|f|) < \infty$ (a.s.), one has the following representation of the Laplace functional transform

$$\mathbb{E}\left[\mathrm{e}^{\tilde{\mu}(f)}\right] = \mathrm{e}^{\psi_\nu(f)} \tag{3}$$

where

$$\psi_\nu(f) = \int_{\mathbb{R}^+ \times \mathbb{X}} \left[1 - \mathrm{e}^{s\, f(x)}\right] \, \nu(\mathrm{d}s, \mathrm{d}x)$$

is also termed the *Laplace exponent* of $\tilde{\mu}$. The main goal we will pursue in the next sections consists in defining a vector of dependent CRMs $(\tilde{\mu}_1, \ldots, \tilde{\mu}_n)$ such that marginally each $\tilde{\mu}_i$ is a gamma CRM. A vector of dependent random probabilities will, then, be obtained by normalizing $\tilde{\mu}_i$, for $i = 1, \ldots, n$. The resulting dependent Dirichlet processes will represent a candidate for modelling partial exchangeable data as in (2) when the cardinality of $\mathcal{Z}$ is n.

3. A multivariate Gamma CRM

The definition outlined in Section 2 can be extended to the case of vectors of CRMs, namely vectors $(\tilde{\mu}_1, \ldots, \tilde{\mu}_n)$ of random elements defined on $(\Omega, \mathcal{F}, \mathbb{P})$ and taking values in $(\boldsymbol{M}^n_{\mathbb{X}}, \mathscr{M}^n_{\mathbb{X}})$ such that for any pair of disjoint sets A_1 and A_2 in $\mathscr{X}$ the vectors $(\tilde{\mu}_1(A_i), \ldots, \tilde{\mu}_n(A_i))$ are, for $i = 1, 2$, independent. In a similar fashion as for the one–dimensional case, a Poisson process type representation holds true in the sense that

$$(\tilde{\mu}_1, \ldots, \tilde{\mu}_n) = \sum_r (J_{r,1}, \ldots, J_{r,n})\, \delta_{X_r}.$$

In the previous representation, $\{(J_{r,1}, \ldots, J_{r,n}, X_r) : r = 1, 2, \ldots\}$ are points from a Poisson process on $((\mathbb{R}^+)^n, \mathbb{X})$ with intensity measure ν such that

$$\int_{(\mathbb{R}^+)^n \times B} \|\boldsymbol{s}\|\, \nu(\mathrm{d}\boldsymbol{s}, \mathrm{d}x) < \infty \qquad \forall B \in \mathscr{X}$$

and with $\nu(A_i \times B) = \nu_i(A \times B)$ for any A in $\mathscr{B}(\mathbb{R}^+)$, where ν_i is the intensity of $\tilde{\mu}_i$ and $A_i = (\mathbb{R}^+)^{i-1} \times A \times (\mathbb{R}^+)^{n-i}$. Moreover, for any collection of measurable functions $f_i : \mathbb{X} \to \mathbb{R}$, $i = 1, \ldots, n$ such that $\tilde{\mu}_i(|f_i|) < \infty$ (a.s.) one has

$$\mathbb{E}\left[\mathrm{e}^{\tilde{\mu}_1(f_1) + \cdots + \tilde{\mu}_n(f_n)}\right] = \mathrm{e}^{-\psi_{\nu,n}(\boldsymbol{f})} \tag{4}$$

where $\boldsymbol{f} = (f_1, \ldots, f_n)$,

$$\psi_{\nu,n}(\boldsymbol{f}) = \int_{(\mathbb{R}^+)^n \times \mathbb{X}} \left[1 - \mathrm{e}^{\langle \boldsymbol{y}, \boldsymbol{f}(x)\rangle}\right]\, \nu(\mathrm{d}\boldsymbol{y}, \mathrm{d}x).$$

and $\langle \boldsymbol{y}, \boldsymbol{f}(x)\rangle = \sum_{i=1}^n y_i f_i(x)$.

Our main goal is the proposal of a vector $(\tilde{\mu}_1, \ldots, \tilde{\mu}_n)$ whose marginals are gamma CRMs. This is accomplished by setting

$$\nu(\mathrm{d}\boldsymbol{y}, \mathrm{d}x) = \sum_{i=0}^{n-1} \frac{(n-1)!}{(n-i-1)!} \frac{e^{-|\boldsymbol{y}|}}{|\boldsymbol{y}|^{i+1}}\, \mathrm{d}\boldsymbol{y}\, H(\mathrm{d}x) \tag{5}$$

with $|\boldsymbol{y}| = \sum_{=1}^n y_i$. It is easy to check that

$$\nu(A_i \times B) = H(B) \int_A \frac{\mathrm{e}^{-y}}{y}\, \mathrm{d}y \qquad i = 1, \ldots, n,$$

for any $A \in \mathscr{B}(\mathbb{R}^+)$ and this clearly implies that $\tilde{\mu}_1, \ldots, \tilde{\mu}_n$ marginally are identically distributed gamma CRMs. It is worth noting that (5) corresponds to the superposition of n vector of CRMs

$$(\tilde{\mu}_1, \ldots, \tilde{\mu}_n) = \sum_{i=1}^n (\mu^*_{i,1}, \ldots, \mu^*_{i,n})$$

with the intensity of the ith summand $(\mu^*_{i,1}, \ldots, \mu^*_{i,n})$ being

$$\nu^*_i(\mathrm{d}\boldsymbol{y}, \mathrm{d}x) = H(\mathrm{d}x)\, \frac{(n-1)!}{(n-i)!} \frac{\mathrm{e}^{-|\boldsymbol{y}|}}{|\boldsymbol{y}|^i}$$

It is interesting to note that $\nu^*_1((\mathbb{R}^+)^n \times B) = \infty$, whereas $\nu^*_i((\mathbb{R}^+)^n \times B) < \infty$ for any $i = 2, \ldots, n$ and for any $B \in \mathscr{X}$. Hence, for any $i \geq 2$ the vector $(\mu^*_{i,n}, \ldots, \mu^*_{n,n})$ has a finite number of jumps and acts as a multivariate Poisson compound process.

The proposal in (5) can also be obtained by applying a specific Lévy copula to the marginal gamma intensities of $\tilde{\mu}_1, \ldots, \tilde{\mu}_n$. Lévy copulas have been recently introduced in [2] as an extension, to Lévy processes, of the familiar notion of copulas for probability distributions. See also [6] and [14] for applications to Bayesian nonparametric inference. The claimed connection is best seen by confining to the case where $n = 2$. We shall use the notation $\Gamma(a, x) = \int_x^\infty s^a \, \mathrm{e}^{-s} \, \mathrm{d}s$ for the incomplete gamma function, whereas $\Gamma^{-1}(a, x)$ is the inverse function of $x \mapsto \Gamma(a, x)$, for any $a \in \mathbb{R}$.

Proposition 1. *The measure ν defined in (5) with $n = 2$ can be recovered by applying the Lévy Copula*

$$C(y_1, y_2) = \Gamma(0, \Gamma^{-1}(0, y_1) + \Gamma^{-1}(0, y_2))$$

to a pair of gamma CRMs.

Proof. The tail integral of each marginal gamma CRM is

$$\mathcal{U}_i(x) = \int_x^{+\infty} y^{-1} e^{-y} \, \mathrm{d}y = \Gamma(0, x) \qquad i = 1, 2$$

On the other hand, the tail integral associated to (5) with $n = 2$ is

$$\mathcal{U}(x_1, x_2) = \int_{x_1}^{+\infty} \int_{x_2}^{+\infty} \left[\frac{1}{(y_1 + y_2)^2} \mathrm{e}^{-y_1 - y_2} + \frac{1}{(y_1 + y_2)} \mathrm{e}^{-y_1 - y_2} \right] \mathrm{d}y_1 \, \mathrm{d}y_2$$

After the change of variable $s = y_1 + y_2$ and $t = y_1$ we obtain

$$\begin{aligned} \mathcal{U}(x_1, x_2) &= \int_{x_1 + x_2}^{+\infty} e^{-s} \left(\frac{1}{s^2} + \frac{1}{s} \right) \int_{x_1}^{s - x_2} \mathrm{d}t \, \mathrm{d}s \\ &= \int_{x_1 + x_2}^{+\infty} e^{-s} \left(\frac{s - (x_1 + x_2)}{s^2} + \frac{s - (x_1 + x_2)}{s} \right) \mathrm{d}s \\ &= \Gamma(0, x_1 + x_2) - (x_1 + x_2)\Gamma(-1, x_1 + x_2) + e^{-x_1 - x_2} \\ &\qquad - (x_1 + x_2)\Gamma(0, x_1 + x_2) \end{aligned}$$

Since $\Gamma(a + 1, x) = a\Gamma(a, x) + x^a \, \mathrm{e}^{-x}$ one has

$$\begin{aligned} \mathcal{U}(x_1, x_2) &= \Gamma(0, x_1 + x_2) - (x_1 + x_2) \left[\frac{e^{-x_1 - x_2}}{x_1 + x_2} - \Gamma(0, x_1 + x_2) \right] \\ &\qquad + e^{-x_1 - x_2} - (x_1 + x_2)\Gamma(0, x_1 + x_2) \\ &= \Gamma(0, x_1 + x_2) \end{aligned}$$

From Theorem 5.3 in Cont and Tankov (2004), the copula C for this process is characterized by

$$\mathcal{U}(x_1, x_2) = C(\mathcal{U}(x_1), \mathcal{U}(x_2))$$

which in this case reduces to

$$\Gamma(0, x_1 + x_2) = C(\Gamma(0, x_1), \Gamma(0, x_2))$$

Setting $y_i = \Gamma(0, x_i)$, for $i = 1, 2$, completes the proof. $\square$

4. The Lévy Exponent

As pointed out in [17], an important tool for possible applications of CRMs to Bayesian nonparametric inference is the Laplace functional transform. This remark is still relevant when working in a multivariate framework and explains why we focus on determining the Lévy exponent induced by (5). We first show in detail how to deal with the case $n = 2$ and, then, deduce an expression for the case for an arbitrary n by induction. Before proceeding it is worth noting that, since $(\tilde{\mu}_1, \ldots, \tilde{\mu}_n)$ has independent increments its distribution is characterized by a choice of $f_1, \ldots, f_n$ in (4) such that $f_i = -\lambda_i \, \mathbb{1}_A$ for any set A in $\mathscr{X}$, $\lambda_i \in \mathbb{R}^+$ and $i = 1, \ldots, n$. In this case

$$\psi_{\nu,n}(\boldsymbol{f}) = H(A)\, \psi^*_{\nu,n}(\boldsymbol{\lambda})$$

where $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)$ and

$$\psi^*_{\nu,n}(\boldsymbol{\lambda}) = \int_{(\mathbb{R}^+)^n} \left[1 - \mathrm{e}^{-\langle \boldsymbol{\lambda}, \boldsymbol{y} \rangle}\right] \sum_{i=0}^{n-1} \frac{(n-1)!}{(n-i-1)!} \frac{e^{-|\boldsymbol{y}|}}{|\boldsymbol{y}|^{i+1}} \, \mathrm{d}\boldsymbol{y} \tag{6}$$

We first deal with the case $n = 2$ since it identifies a structure we can, then, extend to consider any $n > 2$.

Proposition 2. *Let ν be the Lévy intensity introduced in (5) with $n = 2$. The corresponding Lévy exponent has the following form:*

$$\psi^*_{\nu,2}(\lambda_1, \lambda_2) = \begin{cases} [\lambda_1 \log(1+\lambda_1) - \lambda_2 \log(1+\lambda_2)]/(\lambda_1 - \lambda_2) & \lambda_1 \neq \lambda_2 \\ \log(1+\lambda_1) + \lambda_1/(\lambda_1+1) & \lambda_1 = \lambda_2 \end{cases}$$

Proof. Suppose $\lambda_1 \neq \lambda_2$. Correspondingly one has

$$\psi_{\nu,2}(\lambda_1, \lambda_2) = I_1(\lambda_1, \lambda_2) + I_2(\lambda_1, \lambda_2)$$

where

$$I_1(\lambda_1, \lambda_2) = \int_0^\infty \int_0^\infty \left(1 - \mathrm{e}^{-\lambda_1 y_1 - \lambda_2 y_2}\right) \frac{\mathrm{e}^{-y_1 - y_2}}{y_1 + y_2}$$

$$I_2(\lambda_1, \lambda_2) = \int_0^\infty \int_0^\infty \left(1 - \mathrm{e}^{-\lambda_1 y_1 - \lambda_1 y_2}\right) \frac{\mathrm{e}^{-y_1 - y_2}}{(y_1 + y_2)^2}$$

The change of variable $y_1 + y_2 = w$ and $y_1/(y_1+y_2) = z$ leads to

$$
\begin{aligned}
I_1(s,t) &= \int_0^1 \mathrm{d}z \int_0^\infty \left(1 - e^{-w(\lambda_1 z + \lambda_2(1-z))}\right) \mathrm{e}^{-w}\, \mathrm{d}w \\
&= 1 - \frac{\log(1+\lambda_1) - \log(1+\lambda_2)}{\lambda_1 - \lambda_2}
\end{aligned}
$$

and, similarly

$$
\begin{aligned}
I_2(\lambda_1,\lambda_2) &= \int_0^1 \mathrm{d}z \int_0^\infty \left(1 - e^{-w(\lambda_1 z + \lambda_2(1-z))}\right) \frac{e^{-w}}{w}\, \mathrm{d}w \\
&= \frac{1+\lambda_1}{\lambda_1-\lambda_2} \log(1+\lambda_1) - \frac{1+\lambda_2}{\lambda_1-\lambda_2} \log(1+\lambda_2) - 1
\end{aligned}
$$

and combining these two expression one obtains ψ_ν. Proceeding in a similar fashion, and with some useful simplifications, one also obtains $\psi_{\nu,2}(\lambda_1,\lambda_1)$ when $\lambda_1 = \lambda_2$. $\square$

The statement of Proposition 2 points out that one needs to take into account possible ties in the vector $\boldsymbol{\lambda}$ when determining an expression, in closed form, of ψ_ν^*. Hence, when dealing with the case $n > 2$ we shall first assume that $\boldsymbol{\lambda}$ has no ties and, then, move on to the case where any two λ_i and λ_j, with $i \neq j$, may coincide. To this end we need to introduce some further notation. In particular we set $E_n = \{\boldsymbol{x} \in (\mathbb{R}^+)^n : \; x_1 \neq x_2 \neq \cdots \neq x_n\}$ and $\phi_n : E_n \to \mathbb{R}^+$ as

$$
\phi_n(\boldsymbol{x}) = \sum_{i=1}^n \frac{x_i^{n-1} \log(1+x_i)}{\prod_{j=1, j\neq i}^n (x_i - x_j)}. \tag{7}
$$

A preliminary Lemma provides a useful recursive relationship for $\psi_{\nu,n}$

Lemma 1. *Suppose that $\boldsymbol{\lambda} \in E_{n+1}$, for any $n \geq 1$, and denote as $\boldsymbol{\lambda}_{-i}$ the original $\boldsymbol{\lambda}$ vector with the i–th component removed. Then the following recursive equation holds true*

$$
\psi_{\nu,n+1}(\boldsymbol{\lambda}) = \frac{\lambda_{n+1}}{\lambda_{n+1} - \lambda_n} \psi_{\nu,n}(\boldsymbol{\lambda}_{-n}) + \frac{\lambda_n}{\lambda_n - \lambda_{n+1}} \psi_{\nu,n}(\boldsymbol{\lambda}_{-(n+1)}) \tag{8}
$$

Proof. If $A_j^n = \{\boldsymbol{k} \in \{0,1,...,j\}^n : \; |\boldsymbol{k}| = j\}$, then

$$
\begin{aligned}
1 - \mathrm{e}^{-\langle \boldsymbol{\lambda}, \boldsymbol{y} \rangle} &= -\sum_{j\geq 1} \frac{(-1)^j (\langle \boldsymbol{\lambda}, \boldsymbol{y} \rangle)^j}{j!} \\
&= \sum_{j\geq 1} \frac{(-1)^{j+1}}{j!} \sum_{\boldsymbol{k} \in A_j^n} \frac{j!}{k_1! \cdots k_n!} \lambda_1^{k_1} \cdots \lambda_n^{k_n} y_1^{k_1} \cdots y_n^{k_n}
\end{aligned}
$$

and

$$\psi_{\nu,n}(\boldsymbol{\lambda}) = \sum_{i=0}^{n-1} \frac{(n-1)!}{(n-1-i)!} \sum_{j\geq 1} \frac{(-1)^{j+1}}{j!} \sum_{\boldsymbol{k}\in A_j^n} \frac{j!}{k_1!\cdots k_n!} \lambda_1^{k_1} \cdots \lambda_n^{k_n}\ I_n^*(\boldsymbol{k})$$

where

$$I_n^*(\boldsymbol{k}) = \int_{(\mathbb{R}^+)^n} y_1^{k_1}\ \cdots\ y_n^{k_n}\ \frac{\mathrm{e}^{-|\boldsymbol{y}|}}{|\boldsymbol{y}|^{i+1}}\, \mathrm{d}y_1\ \cdots\ \mathrm{d}y_n$$

A simple change of variable, namely $z_i = y_i/s$ for $i = 1, \ldots, n-1$ and $s = |\boldsymbol{y}|$, yields

$$I_n^*(\boldsymbol{k}) = \frac{k_1!\cdots k_n!}{(j+n-1)!}(n-2-i+j)!.$$

This in turn leads to

$$\begin{aligned}
\psi_{\nu,n}(\boldsymbol{\lambda}) &= \sum_{i=0}^{n-1} \frac{(n-1)!}{(n-1-i)!} \sum_{j\geq 1} \sum_{\boldsymbol{k}\in A_j^n} \frac{(-1)^{j+1}\,(n-2-i+j)!}{(j+n-1)!}\ \lambda_1^{k_1}\ \cdots\ \lambda_n^{k_n} \\
&= (n-1)! \sum_{j\geq 1} \sum_{\boldsymbol{k}\in A_j^n} \frac{(-1)^{j+1}}{(j+n-1)!}\ \lambda_1^{k_1}\ \cdots\ \lambda_n^{k_n} \sum_{l=0}^{n-1} \frac{(l+j-1)!}{l!} \\
&= \sum_{j\geq 1} \sum_{\boldsymbol{k}\in A_j^n} \frac{(-1)^{j+1}}{j}\ \lambda_1^{k_1}\ \cdots\ \lambda_n^{k_n} \qquad (9)
\end{aligned}$$

since

$$\sum_{l=0}^{n-1} \frac{(l+j-1)!}{l!} = \frac{1}{j}\frac{(j+n-1)!}{(n-1)!}. \qquad (10)$$

Hence, if one resorts to (9)

$$\psi_{\nu,n+1}(\boldsymbol{\lambda}) = \sum_{j\geq 1} \frac{(-1)^{j+1}}{j}\ \sum_{k_1=0}^{j} \lambda_1^{k_1} \sum_{k_2=0}^{j-k_1} \lambda_2^{k_2}\ \cdots$$

$$\cdots \sum_{k_n=0}^{j-(k_1+\cdots+k_{n-1})} \lambda_n^{k_n} \lambda_{n+1}^{j-(k_1+\cdots+k_n)}$$

Afer some algebra, the last sum above can be rewritten as

$$\sum_{k_n=0}^{j-(k_1+\cdots+k_{n-1})} \lambda_n^{k_n} \lambda_{n+1}^{j-(k_1+\cdots+k_n)} = \lambda_{n+1}^{j-(k_1+\cdots+k_{n-1})} \sum_{k_n=0}^{j-(k_1+\cdots+k_{n-1})} \left(\frac{\lambda_n}{\lambda_{n+1}}\right)^{k_n}$$

$$= \frac{\lambda_{n+1}^{j-(k_1+\cdots+k_{n-1})+1} - \lambda_n^{j-(k_1+\cdots+k_{n-1})+1}}{\lambda_{n+1} - \lambda_n}$$

Hence

$$\psi_{\nu,n+1}(\boldsymbol{\lambda}) = \frac{\lambda_{n+1}}{\lambda_{n+1} - \lambda_n} \sum_{j \geq 1} \frac{(-1)^{j+1}}{j} \sum_{k_1=0}^{j} \lambda_1^{k_1} \sum_{k_2=0}^{j-k_1} \lambda_2^{k_2} \cdots$$

$$\cdots \sum_{k_{n-1}=0}^{j-(k_1+\cdots+k_{n-2})} \lambda_{n-1}^{k_{n-1}} \lambda_{n+1}^{j-(k_1+\cdots+k_{n-1})}$$

$$+ \frac{\lambda_n}{\lambda_n - \lambda_{n+1}} \sum_{j \geq 1} \frac{(-1)^{j+1}}{j} \sum_{k_1=0}^{j} \lambda_1^{k_1} \sum_{k_2=0}^{j-k_1} \lambda_2^{k_2} \cdots$$

$$\cdots \sum_{k_{n-1}=0}^{j-(k_1+\cdots+k_{n-2})} \lambda_{n-1}^{k_{n-1}} \lambda_n^{j-(k_1+\cdots+k_{n-1})}$$

which shows the validity of (8). □

We are now in a position to state and prove the following result

Proposition 3. *For any* $\boldsymbol{\lambda} \in E_n$ *and* $n \geq 1$ *one has*

$$\psi_{\nu,n}(\boldsymbol{\lambda}) = \phi_n(\boldsymbol{\lambda}) \tag{11}$$

Proof. Suppose (11) holds true for n and we shall show that this implies the validity of (11) for $n+1$. By virtue of Proposition 2 the proof is thus completed by induction. Since (11) holds true for n, for any $\boldsymbol{\lambda} \in (\mathbb{R}^+)^{n+1}$ one has

$$\psi_{\nu,n}(\boldsymbol{\lambda}_{-n}) = \frac{\lambda_{n+1}^{n-1} \log(1+\lambda_{n+1})}{\prod_{j=1}^{n-1}(\lambda_{n+1} - \lambda_j)} + \sum_{i=1}^{n-1} \frac{\lambda_i^{n-1} \log(1+\lambda_i)}{(\lambda_i - \lambda_{n+1}) \prod_{j=1, j\neq i}^{n-1}(\lambda_i - \lambda_j)}$$

$$\psi_{\nu,n}(\boldsymbol{\lambda}_{-(n+1)}) = \sum_{i=1}^{n} \frac{\lambda_i^{n-1} \log(1+\lambda_i)}{\prod_{j=1, j\neq i}^{n}(\lambda_i - \lambda_j)}$$

$$= \frac{\lambda_n^{n-1} \log(1+\lambda_n)}{\prod_{j=1}^{n-1}(\lambda_n - \lambda_j)} + \sum_{i=1}^{n-1} \frac{\lambda_i^{n-1} \log(1+\lambda_i)}{(\lambda_i - \lambda_n) \prod_{j=1, j\neq i}^{n-1}(\lambda_i - \lambda_j)}$$

If these two expressions are plugged in the recursive relation (8) one has

$$\psi_{\nu,n+1}(\boldsymbol{\lambda}) = \frac{\lambda_{n+1}^n \log(1+\lambda_{n+1})}{\prod_{j=1}^n (\lambda_{n+1}-\lambda_j)} + \frac{\lambda_n^n \log(1+\lambda_n)}{\prod_{j=1, j\neq n}^{n+1} (\lambda_n - \lambda_j)}$$

$$\sum_{i=1}^{n-1} \left[\frac{\lambda_{n+1}}{\lambda_i - \lambda_{n+1}} - \frac{\lambda_n}{\lambda_i - \lambda_n} \right] \frac{\lambda_i^{n-1} \log(1+\lambda_i)}{(\lambda_{n+1}-\lambda_n) \prod_{j=1, j\neq i}^{n-1} (\lambda_i - \lambda_j)}.$$

After some algebra, one shows that $\psi_{\nu,n+1}$ satisfies (11) and the proof is completed. □

Reasoning in a similar fashion one can get to the following extension of Proposition 3 that takes into account possible ties in $\boldsymbol{\lambda} \in (\mathbb{R}^+)^n$.

Proposition 4. *Let $\boldsymbol{\lambda} \in (\mathbb{R}^+)^n$ be such that it consists of $l \leq n$ distinct values denoted as $\tilde{\lambda}_1, \ldots, \tilde{\lambda}_l$ with respective multiplicities $(n_1, \ldots, n_l)$. Then*

$$\psi_{\nu,n}(\boldsymbol{\lambda}) = \left(\prod_{i=1}^{l} \frac{1}{\Gamma(n_i)} \frac{\partial^{n_i - 1}}{\partial^{n_i-1} \tilde{\lambda}_i} \right) \left(\phi_l(\tilde{\lambda}_1, \ldots, \tilde{\lambda}_l) \prod_{i=1}^{l} \tilde{\lambda}_i^{n_i - 1} \right) \tag{12}$$

Proof. If $B_j = \{i : \ \lambda_i = \tilde{\lambda}_j\}$, for any $j = 1, \ldots, l$, and

$$|\boldsymbol{y}|_j = \sum_{i \in B_j} y_i$$

for any $\boldsymbol{y} \in (\mathbb{R}^+)^n$, one has, similarly to (9),

$$\psi_{\nu,n}(\boldsymbol{\lambda}) = \sum_{i=0}^{n-1} \frac{(n-1)!}{(n-1-i)!} \sum_{j\geq 1} \frac{(-1)^{j+1}}{j!} \sum_{\boldsymbol{k} \in A_j^l} \frac{j!}{k_1! \cdots k_l!} \tilde{\lambda}_1^{k_1} \cdots \tilde{\lambda}_l^{k_l} \, I_n^{**}(\boldsymbol{k})$$

where

$$\begin{aligned} I_n^{**}(\boldsymbol{k}) &= \int_{(\mathbb{R}^+)^n} |\boldsymbol{y}|_1^{k_1} \cdots |\boldsymbol{y}|_l^{k_l} \frac{\mathrm{e}^{-|\boldsymbol{y}|}}{|\boldsymbol{y}|^{i+1}} \, \mathrm{d}y_1 \cdots \mathrm{d}y_n \\ &= \frac{(n-2-i+j)!}{(n+j-1)!} (n_1)_{k_1} \cdots (n_l)_{k_l} \end{aligned}$$

This implies that

$$\psi_{\nu,n}(\boldsymbol{\lambda}) = \sum_{i=0}^{n-1} \frac{(n-1)!}{(n-1-i)!} \sum_{j\geq 1} \frac{(-1)^{j+1}\,(n-2-i+j)!}{(n+j-1)!}$$

$$\times \sum_{\boldsymbol{k}\in A_j^l} \frac{(n_1)_{k_1}\cdots(n_l)_{k_l}}{k_1!\cdots k_l!}\tilde{\lambda}_1^{k_1}\,\cdots\,\tilde{\lambda}_l^{k_l}$$

$$= (n-1)!\sum_{j\geq 1}\frac{(-1)^{j+1}}{(n+j-1)!}\sum_{\boldsymbol{k}\in A_j^l}\frac{(n_1)_{k_1}\cdots(n_l)_{k_l}}{k_1!\cdots k_l!}$$

$$\times\ \tilde{\lambda}_1^{k_1}\cdots\tilde{\lambda}_l^{k_l}\sum_{l=0}^{n-1}\frac{(l+j-1)!}{l!}$$

$$= \sum_{j\geq 1}\frac{(-1)^{j+1}}{j}\sum_{\boldsymbol{k}\in A_j^l}\frac{(n_1)_{k_1}\cdots(n_l)_{k_l}}{k_1!\cdots k_l!}\,\tilde{\lambda}_1^{k_1}\,\cdots\,\tilde{\lambda}_l^{k_l}$$

where the last equality follows from (10). Now note that

$$\frac{(n_i)_{k_i}}{k_i!}\lambda_i^{k_i} = \frac{(k_i+1)_{n_i-1}}{\Gamma(n_i)}\,\tilde{\lambda}_i^{k_i} = \frac{1}{\Gamma(n_i)}\frac{\partial^{n_i-1}}{\partial\tilde{\lambda}_i^{n_i-1}}\,\tilde{\lambda}_i^{n_i-1+k_i}$$

and from this deduce

$$\psi_{\nu,n}(\boldsymbol{\lambda}) = \sum_{j\geq 1}\frac{(-1)^{j+1}}{j}\sum_{\boldsymbol{k}\in A_j^l}\prod_{i=1}^{l}\frac{1}{\Gamma(n_i)}\,\frac{\partial^{n_i-1}}{\partial\tilde{\lambda}_i^{n_i-1}}\,\tilde{\lambda}_i^{n_i-1+k_i}$$

$$= \frac{1}{\prod_{i=1}^{l}\Gamma(n_i)}\,\frac{\partial^{n-l}}{\partial\tilde{\lambda}_1^{n_1-1}\,\cdots\,\partial\tilde{\lambda}_l^{n_l-1}}$$

$$\times\left(\tilde{\lambda}_1^{n_1-1}\cdots\tilde{\lambda}_l^{n_l-1}\sum_{j\geq 1}\frac{(-1)^{j+1}}{j}\sum_{\boldsymbol{k}\in A_j^l}\tilde{\lambda}_1^{k_1}\,\cdots\,\tilde{\lambda}_l^{k_l}\right)$$

which, from (9) and by virtue of the definition of the function ϕ_l in (11), completes the proof of (12). □

5. Distributional properties of a bivariate Dirichlet process

We now focus on the case where $n = 2$ and consider

$$\tilde{p}_i = \frac{\tilde{\mu}_i}{\tilde{\mu}_i(\mathbb{X})} \qquad i = 1, 2 \tag{13}$$

where $(\tilde{\mu}_1, \tilde{\mu}_2)$ is a vector of dependent gamma CRMs whose Lévy intensity is as in (5). It, then, follows that each $\tilde{p}_i$, for $i = 1, 2$, is a Dirichlet process with baseline measure θP_0 and dependence between $\tilde{p}_1$ and $\tilde{p}_2$ is induced by the dependence between $\tilde{\mu}_1$ and $\tilde{\mu}_2$. We shall now point out a few distributional properties of $(\tilde{p}_1, \tilde{p}_2)$ that are of interest for applications to Bayesian nonparametric inference. To this end, let

$$g_\nu(q_1, q_2, s, t) = \int_0^\infty \int_0^\infty y_1^{q_1} y_2^{q_2} e^{-sy_1 - ty_2} \nu(y_1, y_2) \, \mathrm{d}y_1 \, \mathrm{d}y_2 \tag{14}$$

and note that this has been a relevant quantity considered in [14] for characterizing the partition probability function associated to a two–parameter Poisson–Dirichlet process vector. Here we shall make a similar use of g_ν. Before proceeding, however, we provide a closed form expression of (14) when ν is as in (5) with $n = 2$. Note that one can also write

$$g_\nu(q_1, q_2; s, t) = I_1(q_1, q_2; s, t) + I_2(q_1, q_2; s, t)$$

where

$$I_1(q_1, q_2; s, t) = \int_{(\mathbb{R}^+)^2} y_1^{q_1} \, y_2^{q_2} \, \mathrm{e}^{-sy_1 - ty_2} \, \frac{\mathrm{e}^{-y_1 - y_2}}{(y_1 + y_2)^2} \, \mathrm{d}y_1 \, \mathrm{d}y_2$$

$$I_2(q_1, q_2; s, t) = \int_{(\mathbb{R}^+)^2} y_1^{q_1} \, y_2^{q_2} \, \mathrm{e}^{-sy_1 - ty_2} \, \frac{\mathrm{e}^{-y_1 - y_2}}{y_1 + y_2} \, \mathrm{d}y_1 \, \mathrm{d}y_2$$

when $q_1 + q_2 \geq 1$. Morevoer, $g_\nu(0, 0; s, t) \equiv 1$. a simple change of variable into polar coordinates yields

$$\begin{aligned}
I_1(q_1, q_2, s, t) = & \int_0^{\frac{\pi}{2}} \sin(2\theta) \int_{\mathbb{R}^+} \rho^{q_1 + q_2 - 1} \cos^{2q_1}(\theta) \, \sin^{2q_2}(\theta) \\
& \times \, \mathrm{e}^{-\rho[(1+s)\cos^2(\theta) + (1+t)\sin^2(\theta)]} \mathrm{d}\rho \, \mathrm{d}\theta \\
& = \Gamma(q_1 + q_2) \int_0^{\frac{\pi}{2}} \frac{\cos^{2q_1}(\theta) \, \sin^{2q_2}(\theta) \, \sin(2\theta)}{[(1+s)\cos^2(\theta) + (1+t)\sin^2(\theta)]^{q_1 + q_2}} \, \mathrm{d}\theta \\
& = \Gamma(q_1 + q_2) \int_0^1 \frac{y^{q_1}(1-y)^{q_2}}{[(1+s)y + (1+t)(1-y)]^{q_1 + q_2}} \, \mathrm{d}y \\
& = (1+t)^{-q_1 - q_2} \Gamma(q_1 + q_2) \int_0^1 \frac{y^{q_1}(1-y)^{q_2}}{\left[1 - y\frac{t-s}{1+t}\right]^{q_1 + q_2}} \mathrm{d}y \\
& = \frac{\Gamma(q_2 + 1)\Gamma(q_1 + 1)}{(q_1 + q_2)(q_1 + q_2 + 1)} \, \frac{{}_2F_1(q_1 + q_2, q_1 + 1, q_1 + q_2 + 2, \frac{t-s}{1+t})}{(1+t)^{q_1 + q_2}}
\end{aligned}$$

In a similar fashion one determines $I_2(q_1, q_2; s, t)$ thus yielding

$$g_{\nu_2}(q_1, q_2; s, t) = \frac{\Gamma(q_2+1)\Gamma(q_1+1)}{(q_1+q_2+1)(1+t)^{q_1+q_2}} \left\{ \frac{{}_2F_1(q_1+q_2, q_1+1, q_1+q_2+2, \frac{t-s}{1+t})}{q_1+q_2} \right.$$

$$\left. + \frac{{}_2F_1(q_1+q_2+1, q_1+1, q_1+q_2+2, \frac{t-s}{1+t})}{1+t} \right\} \quad (15)$$

The availability of the g_ν function allows us to determine an expression of the mixed moments of the un–normalized vector $(\tilde{\mu}_1(A), \tilde{\mu}_2(A))$, with $A \in \mathscr{X}$. In the sequel, for any two vectors $\boldsymbol{x} = (x_1, \ldots, x_d)$ and $\boldsymbol{y} = (y_1, \ldots, y_d)$ in $\mathbb{N}_0^d$, then $\boldsymbol{x} \prec \boldsymbol{y}$ if either $|\boldsymbol{x}| < |\boldsymbol{y}|$ or $|\boldsymbol{x}| = |\boldsymbol{y}|$ and $x_1 < y_1$ or if $|\boldsymbol{x}| = |\boldsymbol{y}|$ with $x_i = y_i$ for $i = 1, ..., j$ and $x_{j+1} < y_{j+1}$ for some j in $\{1, ..., d\}$.

Proposition 5. *Let $p_j(q_1, q_2, k)$ be the set of vectors $(\boldsymbol{\lambda}, \boldsymbol{s}_1, \ldots, \boldsymbol{s}_j)$ such that the coordinates of $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_j)$ are positive and such that $\sum_{i=1}^j \lambda_i = k$. Moreover, $\boldsymbol{s}_i = (s_{1,i}, s_{2,i})$ are vectors such that $\mathbf{0} \prec \boldsymbol{s}_1 \prec \cdots \prec \boldsymbol{s}_j$ and $\sum_{i=1}^j \lambda_i(s_{1,i} + s_{2,i}) = k = q_1 + q_2$. Then,*

$$\mathbb{E}\left[\prod_{i=1}^{2} \{\tilde{\mu}_i(A)\}^{q_i}\right] = q_1! q_2! \sum_{k=1}^{q_1+q_2} [H(A)]^k \times \sum_{j=1}^{q_1+q_2} \sum_{p_j(q_1,q_2,k)} \prod_{i=1}^{j} \frac{1}{\lambda_i!(s_{1,i}+s_{2,i})^{\lambda_i}}$$

Proof. Hence note that

$$\mathbb{E}\left[\mathrm{e}^{-s\tilde{\mu}_1(A) - t\tilde{\mu}_2(A)} \prod_{i=1}^{2} \{\tilde{\mu}_i(A)\}^{q_i}\right] = (-1)^{q_1+q_2} \frac{\partial^{q_1+q_2}}{\partial s^{q_1}\, \partial t^{q_2}}\, \mathrm{e}^{-H(A)\,\psi(s,t)}$$

and by virtue of Theorem 2.1 in [1] one has that the derivative in the right–hand side above coincides with

$$\mathrm{e}^{-H(A)\psi(s,t)}\, q_1! q_2! \sum_{k=1}^{q_1+q_2} (-1)^k [H(A)]^k \times$$

$$\times \sum_{j=1}^{q_1+q_2} \sum_{p_j(q_1,q_2,k)} \prod_{i=1}^{j} \frac{1}{\lambda_i!(s_{1,i}! s_{2,i}!)^{\lambda_i}} \left(\frac{\partial^{s_{1.i}+s_{2,i}}}{\partial s^{s_{1,i}} \partial t^{s_{2,i}}}\, \psi(s,t)\right)^{\lambda_i}$$

By virtue of the definition of the function g_ν one has

$$\mathrm{e}^{-H(A)\psi(s,t)}\, q_1!q_2! \sum_{k=1}^{q_1+q_2} [H(A)]^k \times \\ \times \sum_{j=1}^{q_1+q_2} \sum_{p_j(q_1,q_2,k)} \prod_{i=1}^{j} \frac{1}{\lambda_i!(s_{1,i}!s_{2,i}!)^{\lambda_i}} \;(g_{\nu_2}(s_{1,i}, s_{2,i}, s, t))^{\lambda_i}\,.$$

Since $\psi(0,0) = 1$ and

$$g_{\nu_2}(s_{1,i}, s_{2,i}, 0, 0) = \frac{s_{1,i}!s_{2,i}!}{s_{1,i} + s_{2,i}}$$

the conclusion follows. $\square$

Proceeding in a similar fashion, in [14] one finds an expression of the covariance between $\tilde{p}_1(A)$ and $\tilde{p}_2(B)$ which is given by

$$\mathrm{Cov}(\tilde{p}_1(A), \tilde{p}_2(B)) = \\ [H(A\cap B) - H(A)H(B)] \int_{(\mathbb{R}^+)^2} \mathrm{e}^{-\psi(s,t)} g_\nu(1,1,s,t)\,\mathrm{d}s\,\mathrm{d}t \quad (16)$$

And for the specific g_ν function displayed in (15), one can obtain a representation of the integral in (16) in terms of mixtures of Gauss hypergeometric functions ${}_2F_1$.

Proposition 6. *Let ν be as in* (5)*, with $n = 2$. Then*

$$\int_{(\mathbb{R}^+)^2} \mathrm{e}^{-\psi(s,t)} g_\nu(1,1,s,t)\,\mathrm{d}s\,\mathrm{d}t = \frac{1}{6}\int_0^1 \frac{z^2}{(1-z)\log(1-z)}\; {}_2F_1(2,2,4,z)\,\mathrm{d}z \\ + \frac{2}{3}\int_0^1 \frac{z}{\log(1-z)} \mathrm{e}^{-\frac{z-1}{z}\log(1-z)}\; {}_2F_1(3,2,4,z)\,\mathrm{d}z$$

Proof. From (15) one has

$$\int_{(\mathbb{R}^+)^2} \mathrm{e}^{-\psi(s,t)} g_\nu(1,1,s,t) dsdt = J_1 + J_2$$

Where

$$J_1 = \frac{1}{6}\int_{(\mathbb{R}^+)^2} \mathrm{e}^{-\psi(s,t)}(1+t)^{-2}\; {}_2F_1(2,2,4,\frac{t-s}{1+t})\,\mathrm{d}s\,\mathrm{d}t$$

$$J_2 = \frac{2}{3}\int_{(\mathbb{R}^+)^2} \mathrm{e}^{-\psi(s,t)}(1+t)^{-3}\; {}_2F_1(3,2,4,\frac{t-s}{1+t})\,\mathrm{d}s\,\mathrm{d}t$$

A simple change of variable yields

$$J_1 = \frac{1}{6}\int_0^1\int_0^1 \mathrm{e}^{w\frac{\log(1-z)}{z}}\mathrm{e}^{-\frac{z-1}{z}\log(1-z)}\,{}_2F_1(2,2,4,z)\,\mathrm{d}w\,\mathrm{d}z$$
$$= \frac{1}{6}\int_0^1 \frac{z^2}{(1-z)\log(1-z)}\,{}_2F_1(2,2,4,z)\,\mathrm{d}z$$

In a similar fashion one determines J_2. □

An interesting application of the vector of Dirichlet random probabilities we have been examining so far concerns Bayesian inference, where the distribution of $(\tilde{p}_1,\ldots,\tilde{p}_k)$ can be used as a prior for modelling partially exchangeable random elements arising from k different populations. This situation is displayed in (2). In particular, when the covariate space $\mathcal{Z}$ consists of two elements, i.e. $M=2$, and set $Q_{\mathcal{Z}}$ as the probability distribution of the vector of dependent Dirichlet processes defined in (13). Since $\tilde{p}_1$ and $\tilde{p}_2$ are, almost surely, discrete there may appear ties within each sample and between the two samples $\boldsymbol{X}^{n_1}(z_1)$ and $\boldsymbol{X}^{n_2}(z_2)$. Hence, the n_1+n_2 data consist of k distinct values forming clusters of sizes $N_1,\ldots,N_k$ Moreover, $N_j = n_{j,1}+n_{j,2}\ge 1$ with $n_{j,1}$ and $n_{j,2}$ denoting the number of observations from $\boldsymbol{X}^{n_1}(z_1)$ and $\boldsymbol{X}^{n_2}(z_2)$, respectively, i the j–th group. If $\boldsymbol{n}_1 = (n_{1,1},\ldots,n_{k,1})$ and $\boldsymbol{n}_2 = (n_{1,2},\ldots,n_{k,2})$ are vectors of non–negative integers in the set

$$\Delta_k(n_1,n_2) := \{(\boldsymbol{n}_1,\boldsymbol{n}_2)\in\mathbb{N}_0^{2k} :\ n_{j,1}+n_{j,2}\ge 1,\ |\boldsymbol{n}_i| = n_i\}$$

we shall denote by

$$\Pi_k^{(n_1,n_2)}(\boldsymbol{n}_1,\boldsymbol{n}_2) = \int_{\mathbb{X}^k}\mathbb{E}\left[\prod_{j=1}^k \tilde{p}_1^{n_{j,1}}(\mathrm{d}x_i)\,\tilde{p}_2^{n_{j,2}}(\mathrm{d}x_i)\right]$$

the probability of detecting two samples $\boldsymbol{X}^{n_1}(z_1)$ and $\boldsymbol{X}^{n_2}(z_2)$ featuring k distinct values with respective frequencies $n_{1,1}+n_{1,2},\ldots,n_{k,1}+n_{k,2}$. Before providing an expression for $\Pi_k^{(n_1,n_2)}$, we need to introduce some notation. In particular, we set

$$\Phi_n(a;z) = \int_0^{1-z} x^n\,\mathrm{e}^{ax}\,\mathrm{d}x$$
$$= \frac{\mathrm{e}^{a(1-z)}}{a^{n+1}}\sum_{i=0}^n(-1)^{n-i}\frac{n!}{i!}\{a(1-z)\}^i + (-1)^{n+1}\frac{n!}{a^{n+1}} \tag{17}$$

and

$$\theta_i(q_1, q_2; z) = \frac{q_1! q_2!}{q_1 + q_2 + 1} \left\{ \frac{{}_2F_1(q_1 + q_2, q_1 + 1; q_1 + q_2 + 2; z)}{q_1 + q_2} \right\}^i \times \left\{ {}_2F_1(q_1 + q_2 + 1, q_1 + 1; q_1 + q_2 + 2; z) \right\}^{1-i} \quad (18)$$

for $i = 0, 1$.

Proposition 7. *For any positive integers n_1, n_2 and k such that $k \leq n_1 + n_2$ and for any $(\boldsymbol{n}_1, \boldsymbol{n}_2) \in \Delta_k(n_1, n_2)$ one has*

$$\Pi_k^{(n_1,n_2)}(\boldsymbol{n}_1, \boldsymbol{n}_2) = \frac{1}{\prod_{i=1}^2 \Gamma(n_i)} \sum_{\boldsymbol{i} \in \{0,1\}^k} \sum_{\ell=0}^{n_1-1} \sum_{m=0}^{n_2-1} \binom{n_1 - 1}{\ell} \binom{n_2 - 1}{m} \times (-1)^{k-|\boldsymbol{i}|+1} \int_0^1 \left\{ \frac{z}{\log(1-z)} \right\}^{\ell+m+k-|\boldsymbol{i}|} (1-z)^{n_1+n_2-2-\ell-m-\frac{z-1}{z}} \times \Phi_{\ell+m+k-|\boldsymbol{i}|}\left(-\frac{\log(1-z)}{z}; z \right) \prod_{j=1}^k \left(\theta_{i_j}(n_{j,1}, n_{j,2}; z) + \theta_{i_j}(n_{j,2}, n_{j,1}; z) \right) \mathrm{d}z \quad (19)$$

where $|\boldsymbol{i}| = i_1 + \cdots + i_k$.

Proof. The result can be deduced from

$$\Pi_k^{(n_1,n_2)}(\boldsymbol{n}_1, \boldsymbol{n}_2) = \frac{1}{\prod_{i=1}^2 \Gamma(n_i)} \left(\int_{A_-} + \int_{A_+} \right) s^{n_1-1} t^{n_2-1} \mathrm{e}^{-\psi(s,t)} \times \prod_{j=1}^k g_\nu(n_{j,1}, n_{j,2}; s, t) \, \mathrm{d}s \, \mathrm{d}t =: I_1 + I_2$$

where the function g_ν is as in (15), $A_- := \{(s,t) \in (\mathbb{R}^+)^2 : s < t\}$ and $A_+ := \{(s,t) \in (\mathbb{R}^+)^2 : s \geq t\}$. We shall explicitly deal with I_1, which is associated to A_-, since an expression for I_2 can be similarly obtained. Resort to the change of variable $z = (t-s)/(1+t)$ and $w = 1/(1+t)$ and note that (z, w) is in the simplex $\mathcal{S}_1 = \{(z,w) \in [0,1]^2 : z + w \leq 1\}$ since $(s,t) \in A_-$.

Hence

$$I_1 = \frac{1}{\prod_{i=1}^2 \Gamma(n_i)} \sum_{|\boldsymbol{i}|\in\{0,1\}^k} \int_{\mathcal{S}_1} w^{k-|\boldsymbol{i}|}(1-z-w)^{n_1-1}(1-w)^{n_2-1}$$

$$\times\ \mathrm{e}^{-\frac{z-1}{z}\log(1-z)-w\frac{\log(1-z)}{z}} \prod_{j=1}^k \Psi_{i_j}(n_{j,1}, n_{j,2}, z)\,\mathrm{d}z\,\mathrm{d}w$$

$$= \frac{1}{\prod_{i=1}^2 \Gamma(n_i)} \sum_{|\boldsymbol{i}|\in\{0,1\}^k} \sum_{\ell=0}^{n_1-1}\sum_{m=0}^{n_2-1} \binom{n_1-1}{\ell}\binom{n_2-1}{m}(-1)^{\ell+m}$$

$$\times \int_{\mathcal{S}_1} w^{\ell+m+k-|\boldsymbol{i}|}(1-z)^{n_1+n_2-2-\ell-m}\mathrm{e}^{-\frac{z-1}{z}\log(1-z)-w\frac{\log(1-z)}{z}}$$

$$\times \prod_{j=1}^k \theta_{i_j}(n_{j,1}, n_{j,2}, z)\,\mathrm{d}z\,\mathrm{d}w$$

and the first part in the representation of $\Pi_k^{(n_1,n_2)}$ follows upon noting that

$$\int_0^{1-z} w^{\ell+m+k-|bmi|}\,\mathrm{e}^{-w\frac{\log(1-z)}{z}}\,\mathrm{d}w = \Phi_{\ell++m+k-|bmi|}\left(-\frac{\log(1-z)}{z};\, z\right)$$

A similar procedure on A_+ leads to an expression of I_2 that completes the proof. □

[1] G.M. Constantines, T.H. Savits (1996). A multivariate version of the Faa di Bruno formula. *Trans. Amer. Math. Soc.* **348**, 503–520.

[2] R. Cont, P. Tankov (2004). *Financial modelling with jump processes.* Chapman & Hall/CRC, Boca Raton, FL.

[3] D.J. Daley, D. Vere-Jones (2003). *An introduction to the theory of point processes. Vol. 1.* Springer, New York.

[4] M. De Iorio, P. Müller, G.L. Rosner, S.N. MacEachern (2004). An ANOVA model for dependent random measures. *J. Amer. Statist. Assoc.* **99**, 205–215.

[5] Dunson, D.B. (2010). Nonparametric Bayes applications to biostatistics. In *Bayesian nonparametrics* (Holmes, C.C., Hjort, N.L., Müller, P. and Walker, S.G., Eds.), 223–273, Cambridge University Press, Cambridge.

[6] I. Epifani, A. Lijoi (2010). Nonparametric priors for vectors of survival functions. *Statistica Sinica* **20**, 1455–1484.

[7] T.S. Ferguson (1973). A Bayesian analysis of some nonparametric problems. *Ann. Statist.* **1**, 209–230.

[8] I. S. Gradshteyn, J.M. Ryzhik (2007). *Table of Integrals, Series, and Products.* 7th Ed. Academic Press, New York.

[9] H. Ishwaran, L.F. James (2001). Gibbs sampling methods for stick-breaking priors. *J. Amer. Stat. Assoc.* **96**, 161–173.

[10] H. Ishwaran, M. Zarepour (2009). Series representations for multivariate generalized gamma processes via a scale invariance principle. *Statist. Sinica* **19**, 1665–1682.

[11] J. Kallsen, P. Tankov (2006). Characterization of dependence of multidimensional Lévy processes using Lèvy copulas. *J. Multivariate Anal.* **97**, 1551–1572.

[12] J.F.C. Kingman (1975). Kingman, J. F. C. (1975). Random discrete distributions (with discussion). *J. Roy. Statist. Soc. Ser. B* **37**, 1–22.

[13] J.F.C. Kingman (1993). *Poisson processes.* Oxford University Press, Oxford.

[14] F. Leisen, A. Lijoi (2011). Vectors of Poisson-Dirichlet processes. *J. Multivariate Anal.* **102**, 482–495.

[15] A. Lijoi, R.H. Mena, I. Prünster (2007). Controlling the reinforcement in Bayesian nonparametric mixture models. *J. Roy. Statist. Soc. Ser. B* **69**, 715–740.

[16] Lijoi, A., Nipoti, B. and Prünster (2011). Bayesian inference with dependent normalized completely random measures. *Carlo Alberto Notebboks* **224**

[17] A. Lijoi, I. Prünster (2010). Beyond the Dirichlet process. *Bayesian Nonparametrics* (Holmes, C.C., Hjort, N.L., Müller, P. and Walker, S.G., Eds.), 80–136, Cambridge University Press, Cambridge.

[18] Lo, A.Y. (1984). On a class of Bayesian nonparametric estimates. I. Density estimates. *Ann. Statist.* **12**, 351–357.

[19] S.N. MacEachern (1999). Dependent nonparametric processes. In *ASA Proceedings of the Section on Bayesian Statistical Science*, Alexandria, VA: American Statistical Association.

[20] J. Pitman (1995). Exchangeable and partially exchangeable random partitions *Probab. Theory Related Fields.* **102**, 145–158.

[21] J. Pitman (1996). Some developments of the Blackwell-MacQueen urn scheme. In *Statistics, Probability and Game Theory* (T.S. Ferguson, L.S. Shapley and J.B. MacQueen, Eds.). IMS Lecture Notes Monogr. Ser., Vol. **30**. Inst. Math. Statist., Hayward, 245–267.

[22] J. Pitman (2006). *Combinatorial Stochastic Processes.* Ecole d'Eté de Probabilités de Saint-Flour XXXII 2002. Lecture Notes in Mathematics **1875**. Springer, Berlin.

[23] Pitman, J. and Yor, M. (1997). The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator. *Ann. Probab.* **25**, 855–900.

[24] E. Regazzini, A. Lijoi, I. Prünster (2003). Distributional results for means of normalized random measures with independent increments. *Ann. Statist.* **31**, 560–585.

[25] E. Regazzini, V.V. Sazonov (2000). Approximation of distributions of random probabilities by mixtures of Dirichlet distributions with applications to nonparametric Bayesian statistical inferences. *Theory Probab. Appl.* **45**, 93–110.

[26] D. Spanò, A. Lijoi (2012). Canonical correlations for dependent gamma processes. *CRiSM Working Paper* **12–12**, University of Warwick, U.K.

[27] Teh, Y.W. and Jordan, M.I. (2010). Hierarchical Bayesian nonparametric models with applications. In *Bayesian nonparametrics* (Holmes, C.C., Hjort, N.L., Müller, P. and Walker, S.G., Eds.), 158–207, Cambridge University Press, Cambridge.